Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7857–7861, November 1988
Biochemistry

A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences

(transcription factors/serum response element)

BARBARA A. CHRISTY, LESTER F. LAU*, AND DANIEL NATHANS

Howard Hughes Medical Institute Laboratory and the Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, MD 21205

Contributed by Daniel Nathans, June 21, 1988

ABSTRACT **We have recently identified by cDNA cloning a set of genes that are rapidly activated in mouse 3T3 cells by serum or purified growth factors. Here we report that the cDNA (clone 268) derived from one of these immediate early genes (*zif/268*) encodes a protein with three tandem "zinc finger" sequences typical of a class of eukaryotic transcription factors. The mRNA of *zif/268* is present in many organs and tissues of the mouse and is especially abundant in the brain and thymus tissue. The 5′ genomic flanking sequence of *zif/268* has sequences related to binding sites for known regulatory proteins, including four sequences that resemble the core of the serum response elements (SREs) upstream of the c-*fos* and actin genes. The SRE-like sequences could be responsible for the coordinate activation of *zif/268* and *fos* after serum stimulation of 3T3 cells.**

Growth factors and other extracellular ligands induce sequential changes in gene expression in target cells (1–4). The initial changes are thought to occur via receptor-mediated activation of messenger systems that modify preexisting transcriptional regulators. Subsequent changes in gene expression are presumably mediated by proteins encoded by genes that are activated in the initial response. We and others have been characterizing genes that are rapidly and transiently activated after stimulation of resting murine 3T3 fibroblastic cells in culture with serum, platelet-derived growth factor (PDGF), or fibroblast growth factor (FGF) (1, 3–7) in the expectation that some of these genes encode transcriptional regulators. Among the "immediate early" genes so far identified that appear to encode such proteins are c-*fos* (8), *junB* (9), c-*jun* (ref. 10; K. Ryder and D.N., unpublished results), *Krox 20* (11), and *fra-1* (12). Here we report that one of the rapidly activated genes previously identified by cDNA cloning (clone 268, ref. 4) encodes a protein with three "zinc finger" sequences (13) characteristic of a class of eukaryotic transcription factors, including TFIIIA (14), Sp1 (15), and SW15 (16). This mouse gene, which we call *zif/268*, is homologous to a gene activated in rat pheochromocytoma cells by nerve growth factor (17). The same gene has also been identified recently by Sukhatme *et al.* (18).

MATERIALS AND METHODS

Cell Culture. BALB/c 3T3 cells were maintained and stimulated with serum as previously described (3). Transfection of NIH 3T3 cells was carried out by using the procedure of van der Eb and Graham (19) followed by glycerol shock (20). Chloramphenicol acetyltransferase (CAT) enzyme activity was assayed as described (21). For measuring serum induction of the *zif/CAT* constructs, duplicate cultures were transfected in medium containing 10% fetal bovine serum and then washed three times with medium containing 0.5% fetal bovine serum. After glycerol shock and Hepes-buffered saline rinse, cells were incubated in medium containing 0.5% fetal bovine serum. At appropriate times prior to harvest (48 hr after transfection), fetal bovine serum was added directly to the medium to a final concentration of 20%.

DNA Sequencing. *zif/268* cDNA isolates were cloned into Bluescript plasmids (Stratagene), and sets of 5′ and 3′ deletions were prepared by using exonuclease III [kindly provided by B. Weiss (Johns Hopkins University) or purchased from Promega Biotec, Madison, WI]. The resulting subclones were sequenced by the dideoxynucleotide chain-termination method (22), using deoxyadenosine 5′-[α-[^{35}S]thio]triphosphate and Sequenase (United States Biochemical, Cleveland).†

Primer Extension and S1 Nuclease Assay. These procedures were carried out as described (9). The primer used was a ^{32}P-labeled oligonucleotide complementary to nucleotides 62–84 of the *zif/268* RNA (Fig. 1).

***In Vitro* Transcription and Translation.** *In vitro* transcription and translation were carried out as described (23), except that T7 RNA polymerase was used for transcription. The cDNA template was the longest cDNA clone (Fig. 1) or a subcloned *Bgl* II restriction fragment containing nucleotides 338–1994.

RNA Blot Hybridization. This procedure was performed as described (2) with 15 μg of total cellular RNA and randomly primed [^{32}P]DNA probes.

***In Situ* Hybridization.** Adult male BALB/c mice were killed by asphyxiation and the brains were fixed and processed for *in situ* hybridization as described (24), except that 5% acetic acid/4% formaldehyde/85% ethanol (vol/vol) was used as a fixative. Hybridization of tissue sections with ^{35}S-labeled sense or antisense RNA probe was carried out as described (25) with slight modification.

RESULTS

Nucleotide Sequence of *zif/268* cDNA. *zif/268* mRNA belongs to a class of mRNAs that we have called immediate early mRNAs (4). It appears within minutes of treatment of quiescent BALB/c 3T3 cells with serum, PDGF, or FGF and is superinduced in the presence of a protein synthesis inhibitor. A clone representing the 3′ end of *zif/268* (clone 268 of ref. 4) was used to isolate a nearly full length cDNA clone from a library prepared from poly(A)$^+$ RNA of BALB/c 3T3

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PDGF, platelet-derived growth factor; FGF, fibroblast growth factor; SRE, serum response element; CAT, chloramphenicol acetyltransferase.

*Present address: Department of Molecular Biology, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611.

†The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04089).

cells 3 hr after stimulation by serum in the presence of cycloheximide. By gel electrophoresis, the size of the *zif/268* mRNA was estimated to be about 3.3 kilobases (kb). The sequence obtained from cDNAs of 2.8 and 3.2 kb is shown in Fig. 1.

The 5′ end of the mRNA was identified by two methods, reverse transcription of the RNA using an oligonucleotide primer and S1 nuclease analysis using a probe made with the same primer. Both procedures resulted in products with

```
                ▼
          CGAGAGATCCCAGCGCGCAGAACTTGGGGAGCCGCCGCCGCGATTCGCCGCCGCCGCCAGCTTCCGCCGCCGCAAGATCGGCCCCTGCCCCAG
      CCCTGCCCCAGCCTCCGCGGCAGCCCTGCGTCCACCACGGGCCGCGGCTACCGCCAGCCTGGGGGCCCACCTACACTCCCCGCAGTGTGCCCCTGCACC
      CCGCATGTAACCCGGCCAACCCCCGGCGAGTGTGCCCTCAGTAGCTTCGGCCCCGGGCTGCGCCCACCACCCAACATCAGTTCTCCAGCTCGCTGGTCC
292   GGG ATG GCA GCG GCC AAG GCC GAG ATG CAA TTG ATG TCT CCG CTG CAG ATC TCT GAC CCG TTC GGC TCC TTT CCT   24
          MET Ala Ala Ala Lys Ala Glu MET Gln Leu MET Ser Pro Leu Gln Ile Ser Asp Pro Phe Gly Ser Phe Pro

367   CAC TCA CCC ACC ATG GAC AAC TAC CCC AAA CTG GAG GAG ATG ATG CTG CTG AGC AAC GGG GCT CCC CAG TTC CTC   49
      His Ser Pro Thr MET Asp Asn Tyr Pro Lys Leu Glu Glu MET MET Leu Leu Ser Asn Gly Ala Pro Gln Phe Leu

442   GGT GCT GCC GGA ACC CCA GAG GGC AGC GGC GGT AAT AGC AGC AGC AGC ACC AGC AGC GGG GGC GGT GGT GGG GGC   74
      Gly Ala Ala Gly Thr Pro Glu Gly Ser Gly Gly Asn Ser Ser Ser Ser Thr Ser Ser Gly Gly Gly Gly Gly Gly

517   GGC AGC AAC AGC GGC AGC AGC GCC TTC AAT CCT CAA GGG GAG CCG AGC GAA CAA CCC TAT GAG CAC CTG ACC ACA   99
      Gly Ser Asn Ser Gly Ser Ser Ala Phe Asn Pro Gln Gly Glu Pro Ser Glu Gln Pro Tyr Glu His Leu Thr Thr

592   GAG TCC TTT TCT GAC ATC GCT CTG AAT AAT GAG AAG GCG ATG GTG GAG ACG AGT TAT CCC AGC CAA ACG ACT CGG  124
      Glu Ser Phe Ser Asp Ile Ala Leu Asn Asn Glu Lys Ala MET Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg

667   TTG CCT CCC ATC ACC TAT ACT GGC CGC TTC TCC CTG GAG CCC GCA CCC AAC AGT GGC AAC ACT TTG TGG CCT GAA  149
      Leu Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn Ser Gly Asn Thr Leu Trp Pro Glu

742   CCC CTT TTC AGC CTA GTC AGT GGC CTC GTG AGC ATG ACC AAT CCT CCG ACC TCT TCA TCC TCG GCG CCT TCT CCA  174
      Pro Leu Phe Ser Leu Val Ser Gly Leu Val Ser MET Thr Asn Pro Pro Thr Ser Ser Ser Ser Ala Pro Ser Pro

817   GCT GCT TCA TCG TCT TCC TCT GCC TCC CAG AGC CCG CCC CTG AGC TGT GCC GTG CCG TCC AAC GAC AGC AGT CCC  199
      Ala Ala Ser Ser Ser Ser Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys Ala Val Pro Ser Asn Asp Ser Ser Pro

892   ATC TAC TCG GCT GCG CCC ACC TTT CCT ACT CCC AAC ACT GAC ATT TTT CCT GAG CCC CAA AGC CAG GCC TTT CCT  224
      Ile Tyr Ser Ala Ala Pro Thr Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala Phe Pro

967   GGC TCG GCA GGC ACA GCC TTG CAG TAC CCG CCT CCT GCC TAC CCT GCC ACC AAA GGT GGT TTC CAG GTT CCC ATG  249
      Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr Pro Ala Thr Lys Gly Gly Phe Gln Val Pro MET

1042  ATC CCT GAC TAT CTG TTT CCA CAA CAA CAG GGA GAC CTG AGC CTG GGC ACC CCA GAC CAG AAG CCC TTC CAG GGT  274
      Ile Pro Asp Tyr Leu Phe Pro Gln Gln Gln Gly Asp Leu Ser Leu Gly Thr Pro Asp Gln Lys Pro Phe Gln Gly

1117  CTG GAG AAC CGT ACC CAG CAG CCT TCG CTC ACT CCA CTA TCC ACT ATT AAA GCC TTC GCC ACT CAG TCG GGC TCC  299
      Leu Glu Asn Arg Thr Gln Gln Pro Ser Leu Thr Pro Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser

1192  CAG GAC TTA AAG GCT CTT AAT ACC ACC TAC CAA TCC CAG CTC ATC AAA CCC AGC CGC ATG CGC AAG TAC CCC AAC  324
      Gln Asp Leu Lys Ala Leu Asn Thr Thr Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg MET Arg Lys Tyr Pro Asn

1267  CGG CCC AGC AAG ACA CCC CCC CAT GAA CGC CCA TAT GCT TGC CCT GTC GAG TCC TGC GAT CGC CGC TTT TCT CGC  349
      Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg

1342  TCG GAT GAG CTT ACC CGC CAT ATC CGC ATC CAC ACA GGC CAG AAG CCC TTC CAG TGT CGA ATC TGC ATG CGT AAC  374
      Ser Asp Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys MET Arg Asn

1417  TTC AGT CGT AGT GAC CAC CTT ACC ACC CAC ATC CGC ACC CAC ACA GGC GAG AAG CCT TTT GCC TGT GAC ATT TGT  399
      Phe Ser Arg Ser Asp His Leu Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys

1492  GGG AGG AAG TTT GCC AGG AGT GAT GAA CGC AAG AGG CAT ACC AAA ATC CAT TTA AGA CAG AAG GAC AAG AAA GCA  424
      Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala

1567  GAC AAA AGT GTG GTG GCC TCC CCG GCT GCC TCT TCA CTC TCT TCT TAC CCA TCC CCA GTG GCT ACC TCC TAC CCA  449
      Asp Lys Ser Val Val Ala Ser Pro Ala Ala Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr Ser Tyr Pro

1642  TCC CCT GCC ACC ACC TCA TTC CCA TCC CCT GTG CCC ACT TCC TAC TCC TCT CCT GGC TCC TCC ACC TAC CCA TCT  474
      Ser Pro Ala Thr Thr Ser Phe Pro Ser Pro Val Pro Thr Ser Tyr Ser Ser Pro Gly Ser Ser Thr Tyr Pro Ser

1717  CCT GCG CAC AGT GGC TTC CCG TCG CCG TCA GTG GCC ACC ACC TTT GCC TCC GTT CCA CCT GCT TTC CCC ACC CAG  499
      Pro Ala His Ser Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Phe Ala Ser Val Pro Pro Ala Phe Pro Thr Gln

1792  GTC AGC AGC TTC CCG TCT GCG GGC GTC AGC AGC TCC TTC AGC ACC TCA ACT GGT CTT TCA GAC ATG ACA GCG ACC  524
      Val Ser Ser Phe Pro Ser Ala Gly Val Ser Ser Ser Phe Ser Thr Ser Thr Gly Leu Ser Asp MET Thr Ala Thr

1867  TTT TCT CCC AGG ACA ATT GAA ATT TGC TAA AGGGAATAAAAGAAAGCAAAGGGAGAGGCAGGAAAGACATAAAAGCACAGGAGGGAAGA  533
      Phe Ser Pro Arg Thr Ile Glu Ile Cys  *
      GATGGCCGCAAGAGGGGCCACCTCTTAGGTCAGATGGAAGATCTCAGAGCCAAGTCCTTCTACTCACGAGTAGAAGGACCGTTGGCCAACAGCCCTTTC
      CACTTACCATCCCTGCCTCCCCCGTCCTGTTCCCCTTTTGACTTCAGCTGCCTGAAACAGCCATGTCCAAGTTCTTCACCTCTATCCAAAGGACTTGAT
      TTGCATGGTATTGGATAAATCATTTCAGTATCCTCTCCATCACATGCCTGGCCCTTGCTCCCTTCAGCGCTAGACCATCAAGTTGGCATAAAGAAAAAA
      AAATGGGTTTGGGCCCTCAGAACCCTGCCCTGCATCTTTGTACAGCATCTGTGCCATGGATTTTGTTTCCTTGGGGTATTCTTGATGTGAAGATAATT
      TGCATACTCTATTGTATTATTTGGAGTTAAATCCTCACTTTGGGGGAGGGGGGAGCAAAGCCAAGCAAACCAATGATGATCCTCTATTTTGTGATGACT
      CTGCTGTGACATTAGGTTTGAAGCATTTTTTTTTTCAAGCAGCAGTCCTAGGTATTAACTGGAGCATGTGTCAGAGTGTTGTTCCGTTAATTTTGTAAA
      TACTGCTCGACTGTAACTCTCACATGTGACAAAGTATGGTTTGTTTGGTTGGGTTTTGTTTTTGAGAATTTTTTTGCCCGTCCCTTTGGTTTCAAAAGT
      TTCACGTCTTGGTGCCTTTTGTGTGACACGCCTTGCCGATGGCTTGACATGCGCAGATGTGAGGGACACGCTCACCTTAGCCTTAAGGGGGTAGGAGTG
      ATGTTTTGGGGGAGGCTTTGAGAGCAAAAACGAGGAAGAGGGCTGAGCTGAGCTTTCGGTCTCCAGAATGTAAGAAGAAAAAATTTAAACAAAAATCTG
      AACTCTCAAAAGTCTATTTTTCTAAACTGAAAATGTAAATTTATACATCTATTCAGGAGTTGGAGTGTTGTGGTTACCTACTGAGTAGGCTGCAGTTTT
      TGTATGTTATGAACATGAAGTTCATTATTTTGTGGTTTTATTTTACTTTGTACTTGTGTTTGCTTAAACAAAGTAACCTGTTTGGCTTATAAACACATT
      GAATGCGCTCTATTGCCCATGGGATATGTGGTGTGTATCCTTCAGAAAAATTAAAAGGAAAAATAAAGAAAAAAAAAAAAAAAA
```

FIG. 1. The cDNA sequence and inferred amino acid sequence of *zif/268*. Numbers appearing in the left margin refer to the first nucleotide listed on that line. Numbers appearing in the right margin refer to the last amino acid encoded on that line. The start of the mRNA is indicated by ▼. Nucleotides 1–12 were determined by sequencing a genomic clone. The portion of the DNA that encodes three potential zinc-finger structures is underlined.

similar mobility, indicating that the 5′ end of the mRNA is about 11 nucleotides upstream of the longest cDNA clone. This site is 25 nucleotides downstream from a TATA-like element, AAATAGA (see below).

There are 5 ATG codons within the first 400 nucleotides of cDNA sequence, at positions 197, 295, 316, 325, and 379. The first ATG, at position 197, is followed immediately by a termination codon. The 2nd, 3rd, 4th, and 5th ATG codons are all in the same long open reading frame; there are no in-frame termination codons upstream of these ATGs. Although the ATG codon at position 379 is in the most favorable context for initiation (26), *in vitro* translation of *zif/268* transcripts indicates that one of the preceding AUGs is the preferred initiation codon in a reticulocyte extract. As shown in Fig. 2, a nearly full-length transcript of *zif/268* cDNA directs the synthesis of a protein of electrophoretic mobility lower than that directed by a truncated transcript missing the three upstream AUGs. The first of the upstream in-frame ATGs (at position 295) is followed by an open reading frame of 1599 nucleotides, a termination codon, 1223 nucleotides of nontranslated 3′ sequence, and a 3′ poly(A) tail. Polyadenylylation signal sequences (27, 28) are found at positions 1901, 3094, and 3107. Between these polyadenylylation signals are two ATTTA sequences; multiple corresponding sequences of this type in the 3′ untranslated region of a mRNA have been shown to be related to message instability (29), which is a general property of immediate early mRNAs (4).

Predicted Protein Sequence Derived from the *zif/268* cDNA. Fig. 1 shows the amino acid sequence of the protein encoded by the long open reading frame of the *zif/268* cDNA. If translation begins at ATG 295, the predicted protein would be 533 amino acids in length. The most notable feature of the amino acid sequences is the presence of spaced cysteine and histidine residues near the carboxyl end of the protein that correspond to three tandem copies of the consensus sequence for a postulated DNA- and metal-binding domain, the zinc finger structure (13). Compared to published sequences of putative zinc finger-containing proteins, the *zif/268* product is nearly identical to NGFI-A, the product of a gene activated by nerve growth factor in rat PC12 pheochromocytoma cells (17). The *zif/268* product has 18 amino acid substitutions or deletions relative to NGFI-A.

Serum Responsiveness of a *zif/CAT* Plasmid. To determine whether nucleotide sequences upstream of the *zif/268* gene confer serum responsiveness on its promoter, as found for c-*fos* (30, 31), a cloned fragment of *zif/268* genomic DNA containing 338 nucleotides of 5′ cDNA sequence and about 2.5 kb of contiguous upstream sequence was ligated to the coding region of the bacterial *CAT* gene (21). This test plasmid (*zif/CAT*) was transfected into NIH/3T3 cells and the time course of expression of CAT activity was measured after serum starvation and stimulation of the cells with 20% fetal bovine serum. As shown in Fig. 3, the *zif/CAT* construct was induced by serum. The overall level of induction was about 8.3-fold, and the peak of enzyme induction was at about 2–4 hr. The *zif/268* promoter had no detectable activity when the genomic fragment was linked to the *CAT* gene in the opposite orientation. When a *zif/CAT* construct containing only 1 kb of 5′ genomic sequence was used, serum responsiveness was also observed. We conclude that the *zif/268* promoter and sequences that confer serum responsiveness are contained within 1 kb of the start site of transcription.

Nucleotide Sequence of the *zif/268* Upstream Sequences. The sequence of the *zif/268* promoter region sufficient to confer serum responsiveness is shown in Fig. 4. The previously noted sequence AAATAGA at −25 relative to the transcription start site presumably functions as a TATA-like element. Also present are two CCAAT sequences (36) at positions −227 and −336 and core Sp1 sequences (37). Upstream of the TATA-like element of *zif/268* are four repeats of a sequence resembling or identical to the core of the serum response element (SRE) found 5′ of the c-*fos* gene and shown to mediate serum responsiveness (30, 38). However, none of the four *zif/268* sequences shows the degree of symmetry outside the core sequence seen in the c-*fos* or *Xenopus* actin SREs (39). It remains to be seen whether one or more of the *zif/268* sequences actually functions as an SRE.

Presence of *zif/268* mRNA in Mouse Cell Lines and Tissues. To determine how widely expressed the *zif/268* gene is in mouse cell lines and tissues, total RNA from various proliferating lines and from adult tissues were analyzed by blot hybridization analysis for the presence of *zif/268* RNA. The RNA was detected in many cell lines, including C127 cells, I-7 cells, Friend leukemia cells, NB41 neuroblastoma cells, and various teratocarcinoma cell lines. Of the tissues exam-

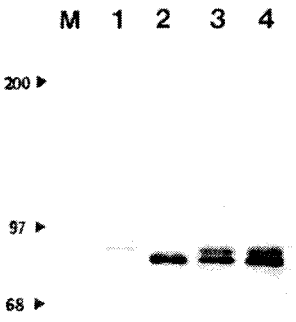

FIG. 2. *In vitro* translation of *zif/268* transcripts. RNA transcripts from a nearly full-length cDNA clone or from a truncated clone containing a *Bgl* II cDNA fragment (nucleotides 338–1994) were translated in a rabbit reticulocyte extract in the presence of [^{35}S]methionine. The products were analyzed by electrophoresis in a sodium dodecyl sulfate/10% acrylamide gel. Lane M, molecular weight markers (given × 10^{-3}); lane 1, products of large transcript; lane 2, products of truncated transcript; lane 3, products of both transcripts cotranslated; lane 4, products 1 and 2 mixed; lane 5, no transcript. Note that *zif/268* product has an anomalously low electrophoretic mobility.

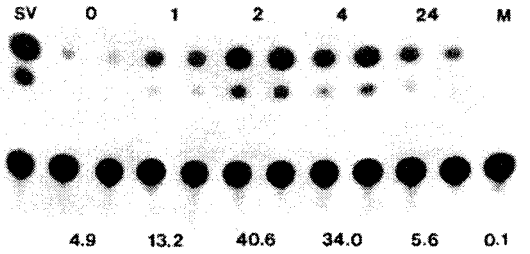

FIG. 3. Serum responsiveness of the *zif/268* promoter in NIH 3T3 cells. Cells were transfected in duplicate with the *zif/CAT* plasmid and treated as described in the text. The figure is an autoradiogram of [^{14}C]chloramphenicol and its acetylated products after chromatography. Numbers at the bottom of the figure are average percent conversion (per 100 μg of total protein) of chloramphenicol to acetylated form in 1 hr; lane SV, cells transfected with simian virus 40–CAT plasmid, pSV2-CAT (21); lane M, mock-transfected; numbers at the top, hr of stimulation with 20% serum prior to harvest. CAT activity of transfected cells in the absence of added serum varied by less than 1.5-fold during the 24-hr incubation.

```
AACAGATCCT GGCGGGGACT TAGGACTGAC CTAGAACAAT CAGGGTTCCG CAATCCAGGT CCCAAAGGTG GGATCCTCAA CCGCAGGACG GAGGGAATAG CCTTTCGATT CTGGGTGGTG
-1023     AP2-like
CATTGGAAGC CCCAGGCTCT AAAACCCCCA ACCTACTGAC TGGTGGCCGA GTATGCACCC GACTGCTAGC TAGGCAGTGT CCCAAGAACC AGTAGCCAAA TGTCTTGGCC TCAGTTTTCC
-903
CGGTGACACC TGGAAAGTGA CCCTGCCATT AGTAGAGGCT CAGGTCAGGG CCCCGCCTCT CCTGGGCGGC CTCTGCCCTA GCCCGCCCTG CCGCTCCTCC TCTCCGCAGG CTCGCTCCCA
-783                                                  AP1-like
CGGTCCCCGA GGTGGGCGGG TGAGCCCAGG ATGACGGCTG TAGAACCCCG GCCTGACTCG CCCTCGCCCC CGCGCCGGGC CTGGGCTTCC CTAGCCCAGC TCGCACCCGG GGGCCGTCGG
-663
AGCCGCCGCG CGCCCAGCTC TACCGGCCTG GCGCCCTCCC CACGCGGGCG TCCCCGACTC CCGCGCGCGC TCAGGCTCCC AGTTGGGAAC CAAGGAGGGG GAGGATGGGG GGGGGGGTGT
-543          SRE-like                                SRE-like
GCGCCGACCC GGAAACGCCA TATAAGGAGC AGGAAGGATC CCCCGCCGGA ACAGGCCTTA TTTCCCCAGC GCCTTATATG GAGTGGCCCA ATATGGCCCT GCCGCTTCCG GCTCTGGGAG
-423
GAGGGGCGAG CGGGGGTTGG GGCGGGGGCA AGCTGGGAAC TCCAGGCGCC TGGCCCGGGA GGCCACTGCT GCTGTTCCAA TACTAGGCTT TCCAGGAGCC TGAGCGCTCG CGATGCCGGA
-303                                   CRE-like                                SRE-like                SRE-like
GCGGGTCGCA GGGTGGAGGT GCCCACCACT CTTGGATGGG AGGGCTTCAC GTCACTCCGG GTCCTCCCGG CCGGTCCTTC CATATTAGGG CTTCCTGCTT CCCATATATG GCCATGTACG
-183
TCACGGCGGA GGCGGGCCCG TGCTGTTCCA GACCCTTGAA ATAGAGGCCG ATTCGGGGAG TCGCGAGAGA TCCCAGCGCG CAGAA
-63                                                                                                                    +1
```

FIG. 4. Nucleotide sequence of the *zif/268* 5′ flanking region. Numbers on the left side refer to the first nucleotide listed on the line directly above, with mRNA start site designated as +1. The TATA-like element, CCAAT sequences, and GC boxes (which contain the core sequences of Sp1 binding sites) are underlined. Boxed sequences: a sequence identical to an activator protein 2 (AP-2) site found in the human growth hormone promoter (32) at position −887 to −894; a sequence resembling an AP-1 binding site (33, 34) at position −604 to −610; a sequence resembling a cAMP-responsive element (CRE) (35) at position −130 to −137; four sequences that resemble the core sequences of the serum-response elements (SREs) of the c-*fos* and actin genes at positions −397 to −406, −343 to −352, −95 to −104, and −73 to −82. Additional sequence (to nucleotide −1383) is included in the sequence data being deposited in the EMBL/GenBank data base.

ined, brain and thymus had a high level of *zif/268* mRNA, while lung and heart had moderate levels (Fig. 5). Low but detectable levels were found in almost all other tissues tested, with the exception of liver. (However, *zif/268* mRNA was found in regenerating liver within 1 hr after partial hepatectomy.) The RNA was also present in mouse embryos and in placenta. In those tissues in which *zif/268* mRNA is most abundant, lesser amounts of smaller hybridizing RNAs were found, including an RNA of about 1.6 kb, a size consistent with utilization of the poly(A) signal at nucleotide 1901 (Fig. 1). To localize *zif/268* mRNA within the brain, we analyzed its distribution in brain sections by *in situ* hybridization (Fig. 6). The RNA is most readily detected in the cerebral cortex and hippocampus.

DISCUSSION

The main conclusion of this report is that the immediate early growth response gene (clone 268), previously identified by cDNA cloning from mRNA of BALB/c 3T3 cells stimulated with serum (4), encodes a protein of 533 amino acids with three typical zinc finger sequences characteristic of a class of eukaryotic transcription factors (40). *zif/268* is homologous to a gene (*NGFI-A*) activated in rat pheochromocytoma cells by nerve growth factor (17); the same gene has been independently identified by Sukahtme *et al.* (18) and designated *Egr-1*. Within the zinc finger region *zif/268* is closely related to the recently described *Krox* gene family (47) and more distantly related to *Sp1* (15). Outside the zinc finger sequences, however, the sequence of *zif/268* is quite different from the sequences of *Krox 20* [the one member of the *Krox* family whose complete cDNA sequence has been reported (11)] and *Sp1* (15).

In addition to its expression in mouse cell lines stimulated with serum or growth factors, *zif/268* is expressed in a number of mouse organs and tissues, notably brain, thymus, lung, and heart (Fig. 5 and ref. 18). In mouse brain *zif/268* mRNA is especially prominent in the cortex and hippocampus, as visualized by *in situ* hybridization (Fig. 6). In more extensive studies of the distribution of *zif/268* mRNA in rat brain, Saffen *et al.* (41) have found that administration of convulsants causes a prompt increase in the mRNA in neurons of the hippocampus and the entorhinal cortex, suggesting that *zif/268* plays a role in regulating the genomic response to stimulation of nerve cells.

The activation of the *zif/268* gene in 3T3 cells by serum, PDGF, or FGF is coordinate with that of c-*fos* (4, 5). The inducibility of c-*fos* expression by serum is mediated by a 22-base-pair sequence with dyad symmetry 5′ to the *fos*

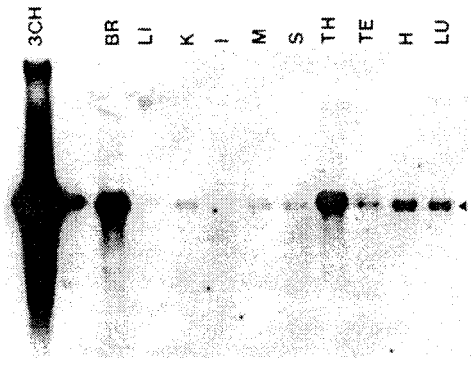

FIG. 5. Blot hybridization analysis of mouse tissue RNA. Total RNA (15 μg) from various mouse tissues was fractionated by electrophoresis and blotted onto nitrocellulose, and the filter was probed with a nearly full-length cDNA clone for *zif/268*. The lane marked 3CH contains 1 μg of total RNA from BALB/c 3T3 cells stimulated with serum for 3 hr in the presence of cycloheximide. Other abbreviations: BR, brain; LI, liver; K, kidney; I, intestine; M, muscle; S, spleen; TH, thymus; TE, testes; H, heart; and LU, lung.

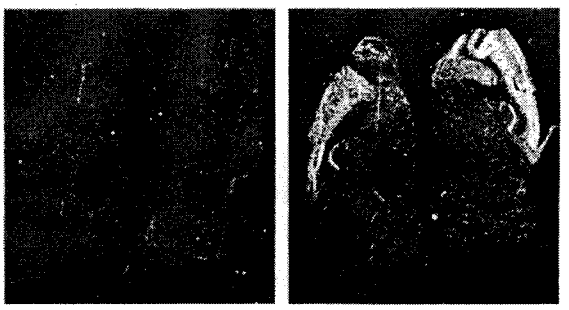

FIG. 6. Localization of *zif/268* mRNA in mouse brain by *in situ* hybridization. Adult mouse brain was fixed and sectioned sagitally. Hybridizations were performed as described in the text, using either the antisense probe (*Right*) or the sense probe (*Left*) for the *zif/268* cDNA. Photographs shown are negative images of film exposed on the slides for 3 days. The regions displaying the highest levels of *zif/268* hybridization were the hippocampus and the cerebral cortex.

TATA signal (30, 31). This SRE is the binding site of a cellular protein thought to serve as a transcriptional activator of *fos* (38, 39, 42) and perhaps other genes coordinately regulated with *fos* (39). It is therefore of interest that upstream of the *zif/268* TATA-like sequence there are four possible SREs. All four share the common core motif CC(A or T)$_6$GG, one of which is identical in sequence to that of *fos*, and another is identical to an artificial SRE (42), but they all lack the extensive flanking sequence symmetry seen in the c-*fos* and other active SREs (39, 43). Whether one or more of these putative SREs is functional remains to be determined. If *zif/268* does have multiple functional SREs, this could explain the brisk response of this gene to serum and growth factors (4, 5, 17). In addition to potential SREs *zif/268* has other possible transcription factor binding sites, namely core Sp1-like sites (37), CCAAT sequences (36), an AP1-like site (33, 34), an AP-2 site (32, 44), and a site similar to the consensus sequence for cAMP-regulated promoters (35) (Fig. 4).

The emerging picture of the early genomic response to growth factors is one of considerable complexity. A large number of genes are rapidly activated in 3T3 cells as part of the direct response to receptor-mediated second messenger systems. Several of these immediate early genes encode known or probable transcription factors: c-*fos* (8), *junB* (9), c-*jun* (AP1) (33, 34), *Krox 20* (11), and *zif/268* (*Egr-1*, *NGFI-A*). Others encode secreted proteins (7) or membrane proteins (S. Hartzell, K. Ryder, and D.N., unpublished results). A possible role of immediate early transcription factors is the repression of immediate early genes. Another possible role is the activation of genes expressed later in the growth response, including those whose products regulate DNA replication. At either site of action, the transcription factors may act independently or as part of a multicomponent transcription complex (10, 45, 46). The first step in elucidating the possible role of the *zif/268*-encoded protein in gene regulation is to find genes with binding sites for this protein.

Note Added in Proof. The nucleotide sequence of a cDNA essentially identical to that of *zif/268* has recently been reported by Lemaire *et al.* (48).

We thank Se-Jin Lee for the embryonic, placental, and cell line RNAs; Stephen Desiderio and Keith Peden for critically reviewing the manuscript; Laura Sanders and Michael McLane for technical assistance; Clark Riley, Lawrence Cisek, and Jeffry Corden for computer assistance; and Lily Mitchell for preparing the manuscript. This research was supported in part by Grant 5 P01 CA16519 from the National Cancer Institute. L.F.L. was a Postdoctoral Fellow of the Helen Hay Whitney Foundation.

1. Cochran, B. H., Reffel, A. C. & Stiles, C. D. (1983) *Cell* **33**, 939–947.
2. Linzer, D. I. H. & Nathans, D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4271–4275.
3. Lau, L. F. & Nathans, D. (1985) *EMBO J.* **4**, 3145–3151.
4. Lau, L. F. & Nathans, D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1182–1186.
5. Sukhatme, V. P., Kartha, S., Toback, F. G., Taub, R., Hoover, R. G. & Tsai-Morris, C.-H. (1987) *Oncogene Res.* **1**, 343–355.
6. Almendral, J. M., Sommer, D., Macdonald-Bravo, H., Burckhardt, J., Perera, J. & Bravo, R. (1988) *Mol. Cell. Biol.* **8**, 2140–2148.
7. Rollins, B. J., Morrison, E. D. & Stiles, C. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3738–3742.
8. Greenberg, M. G. & Ziff, E. B. (1984) *Nature (London)* **311**, 433–438.
9. Ryder, K., Lau, L. F. & Nathans, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1487–1491.
10. Rauscher, F. J., III, Cohen, D. R., Curran, T., Bos, T. J., Vogt, P. K., Bohmann, D., Tjian, R. & Franza, B. R., Jr. (1988) *Science* **240**, 1010–1016.
11. Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R. & Charnay, P. (1988) *EMBO J.* **7**, 29–35.
12. Cohen, D. R. & Curran, T. (1988) *Mol. Cell. Biol.* **8**, 2063–2069.
13. Miller, J., McLachlan, A. D. & Klug, A. (1985) *EMBO J.* **4**, 1609–1614.
14. Vrana, K. E., Churchill, M. E. A., Tullius, T. D. & Brown, D. C. (1988) *Mol. Cell. Biol.* **8**, 1684–1696.
15. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987) *Cell* **51**, 1079–1090.
16. Stillman, D. J., Bankier, A., Seddon, A., Gorehout, G. & Nasmyth, K. A. (1988) *EMBO J.* **7**, 485–494.
17. Milbrandt, J. (1987) *Science* **238**, 797–799.
18. Sukhatme, V. P., Cao, X., Chang, L. C., Tsai-Morris, C.-H., Stamenkovitch, D., Ferreira, P. C. P., Cohen, D. R., Edwards, S. A., Shows, T. B., Curran, T., LeBeau, M. M. & Adamson, E. D. (1988) *Cell* **53**, 37–43.
19. van der Eb, A. J. & Graham, F. L. (1980) *Methods Enzymol.* **65**, 826–839.
20. Frost, E. & Williams, J. (1978) *Virology* **91**, 39–50.
21. Gorman, C., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1051.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Pelletier, J. & Sonnenberg, N. (1985) *Cell* **40**, 515–526.
24. Lee, S.-J., Talmantes, F., Wilder, E., Linzer, D. I. H. & Nathans, D. (1988) *Endocrinology* **122**, 1761–1768.
25. Segal, D. & Shiloh, B.-Z. (1986) *Mol. Cell. Biol.* **6**, 2241–2248.
26. Kozak, M. (1987) *J. Mol. Biol.* **196**, 947–950.
27. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
28. Wickens, M. & Stephenson, P. (1984) *Science* **226**, 1045–1051.
29. Shaw, G. & Kamen, R. (1986) *Cell* **46**, 659–667.
30. Treisman, R. (1985) *Cell* **42**, 889–902.
31. Greenberg, M. E., Siegried, Z. & Ziff, E. B. (1987) *Mol. Cell. Biol.* **7**, 1217–1225.
32. Imagawa, M., Chiu, R. & Karin, M. (1987) *Cell* **51**, 251–260.
33. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P. & Karin, M. (1987) *Cell* **49**, 729–739.
34. Lee, W., Mitchell, P. & Tjian, R. (1987) *Cell* **49**, 741–752.
35. Montminy, M. R., Sevarino, K. A., Wagner, J. A., Mandel, G. & Goodman, R. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6682–6686.
36. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
37. Kadonaga, J. T., Jones, K. A. & Tjian, R. (1986) *Trends Biochem. Sci.* **11**, 20–23.
38. Treisman, R. (1986) *Cell* **46**, 567–574.
39. Mohun, T., Garrett, N. & Treisman, R. (1987) *EMBO J.* **6**, 667–673.
40. Klug, A. & Rhodes, D. (1987) *Trends Biochem. Sci.* **12**, 464–469.
41. Saffen, D. W., Cole, A. J., Worley, P. F., Christy, B. A., Ryder, K. & Baraban, J. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7795–7799.
42. Treisman, R. (1987) *EMBO J.* **6**, 2711–2717.
43. Prywes, R. & Roeder, R. G. (1987) *Mol. Cell. Biol.* **7**, 3482–3489.
44. Mitchell, P. J., Wang, C. & Tjian, R. (1987) *Cell* **50**, 847–861.
45. Rauscher, F. J., III, Sambucetti, L. C., Curran, T., Distel, R. J. & Spiegelman, B. M. (1988) *Cell* **52**, 471–480.
46. Franza, B. R., Jr., Rauscher, F. J., III, Josephs, S. F. & Curran, T. (1988) *Science* **239**, 1150–1153.
47. Chavrier, P., Lemaire, P., Relevant, O., Bravo, R. & Charnay, P. (1988) *Mol. Cell. Biol.* **8**, 1319–1326.
48. Lemaire, P., Revelant, O., Bravo, R. & Charnay, P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4691–4695.